Appendix D - MIME and Mapping

What is MIME?

"MIME" stands for Multipurpose Internet Mail Extensions. MIME serves two major purposes – it allows mail applications to tell one another what sort of data is in mail, and it also provides standard ways for mail applications to encode data so that it can be sent through the Internet mail system.

MIME Encoding

The Internet uses the "SMTP" protocol to move mail around. SMTP is limited to the US-ASCII character set (see Appendix C). This is a problem for people who speak languages other than American English and so need accented characters or non-American letters, or for people who want to use special symbols like the bullet.

MIME provides a way around this restriction. It offers two encodings, "quoted-printable" and "base64." These encodings use US-ASCII character codes to represent any sort of data you like, including special characters or even non-text data.

Quoted-printable is used for data that is mostly text, but has special characters or very long lines. Quoted-printable looks just like regular text, except when a special character is used. The special character is replaced with an "=" and two more characters that represent the character code of the special character. So, a bullet in quoted-printable looks like "=95."

However, there are some other things that quoted-printable does. For one, since it uses an "=" to mean something special, equal signs must themselves be encoded (as "=3D"). Second, no line in quoted-printable is allowed to be more than 76 characters long. If your mail has a line longer than 76 characters, the quoted-printable encoding will break your line in two, and put an "=" at the end of the first line, to signal to the mail reader at the other end that the two lines are really supposed to be one. Finally, a few mail systems either add or remove spaces from the ends of lines. So, in quoted-printable, any space at the end of a line gets encoded (as "=20") to protect it from such mail systems.

Let's try an example. Here's a passage of text that you might type on your PC:

```
«Il est démontré, disait-il, que les choses ne peuvent être autrement; car tout
étant fait pour une fin, tout est necessairement pour la meilleure fin.»
```

Without any encoding, this might show up on your recipient's screen as:

```
+Il est dimontri, disait-il, que les choses ne peuvent btre autrement; car tout
ıtant fait pour une fin, tout est nicessairement pour la meilleure fin.;
```

This corruption happens because SMTPcannothandle the special characters. However, ifyou and yourrecipient both have MIME, quoted-printable encoding would be used, and yourtext would show up properly:

```
«Il est démontré, disait-il, que les choses ne peuvent être autrement; car tout
étant fait pour une fin, tout est nécessairement pour la meilleure fin.»
```

While your mail was actually in transit, however, it would have looked like:

```
=ABIl est d=E9montr=E9, disait-il, que les choses ne peuvent =EAtre =
autrement; car tout =E9tant fait pour une fin, tout est n=E9cessairement =
pour la meilleure fin.=BB
```

Base64 encoding is another way to protect binary data from the SMTP mail system. However, Base64 makesno attempt to be legible, andis most appropriate for non-text data.

MIME Labelling

The other important part of MIME is that it lets mailers communicate what kind of data is in a message (or part ofa message). The primary mechanism used for this is the Content-Type header:

```
Content-Type: text/plain; charset=iso-8859-1
```

A content-type header is divided into three parts; the content type, the content subtype, and the parameters. In this case, the content type is ''text,'' meaning the message contains mostly legible text. The content subtype is ''plain," which means there aren 't any formatting commandsor anything like that embeddedin the text. Finally, "charset=iso-8859-1" is a parameter; in this case, it identifies the character set the message uses.

The major content types are:

text	legible text
image	pictures and graphics
audio	sound
video	moving pictures
message	messages or pieces of messages
multipart	several different kinds of data in a single message

Practical Issues

There are really only two things you sometimes need to do with Eudora and MIME. One is that it may occasionally be necessary to turn off quoted-printable encoding. Another is that you may want to know how to define mappings between PC file extensions, MIME types, and Macintosh types.

Turning Off Quoted-Printable Encoding

Eudora automatically uses quoted-printable encoding if your mail contains special characters. Eudora also uses quoted-printable encoding for attached plain text files. If your recipients don't have MIME, quoted-printable may hurt more than it helps. If that's the case, just turn off the QP button in the message Toolbar when you are sending text files to those recipients.

Mapping Between File Extensions, MIME Types, and Macintosh Types

Since Eudora needs to have the appropriate extensions on attachment filenames in order to open them up from the message, Eudora has the ability to map between file extensions, MIME types and subtypes, and Macintosh creators and types. Messages received by Eudora can grab the MIME type/subtype and/or Macintosh creator/type from an attachment and map that into the correct file extension. Also, on outgoing messages, Eudora can make sure that attachments are encoded with the correct MIME type/subtype and/or Macintosh creator/type depending on the file extension of the attachment being sent.

Eudora knows about some MIME types. However, since new MIME types are being defined all the time, it maybe necessary to add to Eudora's knowledge from time to time. Adding new mappings between the various types only requires editing the EUDORA.INI file with a text editor (like the one that comes with Eudora).

There is a section in the EUDORA.INI file labelled [Mappings], followed by some entries, one per line. Each entry is called a map. A map defines when the mapping should occur (which can be "in," "out," or "both"), followed by an equal sign and five parameters. These five parameters are (in order) the PC file extension, the Macintosh creator code, the Macintosh type, the MIME type, and the MIME subtype. Here are some sample entries:

```
[Mappings]
both=gif, , , image, gif
both=mpg, , , video, mpeg
both=doc , MSWD, , ,
in=xls, XCEL, , ,
out=xls, XCEL, XLS4, ,
both=eps, , EPSF, application, postscript
```

A map marked "in" only tries to match the map to messages that you receive. A map marked "out" only tries to match the map to messages that you send. A map marked "both" tries to match the map to both incoming and outgoing messages.

The first map above says that any incoming MIME message that has a part type of "image" and subtype of "gif" will get saved to a file with the extension ".gif." It also specifies that outgoing messages that have an attachment with the file extension ".gif" will get the MIME type of "image" and subtype of "gif" if the encoding method of the message is MIME. The second map is similar to the first map in structure, but uses a different file extension and MIME type and subtype.

You can use map entries to move between PC file extensions and Macintosh " creator and type as well. The third map says that if an incoming message has an attachment with the Macintosh creator "MSWD" (which is the Macintosh creator for Microsoft Word) then the file extension of the attachment when saved to disk should be ".dot" (the file extension that Word for Windows uses). Since the map is marked as "both," it will also give attachments with the extension ".dot" on outgoing messages the Macintosh creator of "MSWD" if the encoding method of the message is BinHex.

Note that the Macintosh type from this map is empty. This allows multiple types to be recognized with just one mapping. This is nice for "in" maps because it allows you to cover a range of creator/type pairs with one map. You must be careful in using this type of map with an "out" or "both" mapping. though, because an outgoing attachment that matched this map would have a Macintosh creator, but no Macintosh type. Some Macintosh applications cannot open files with a missing type. Microsoft Word for the Macintosh can open files without a type, so this map is fine being marked "both."

Microsoft Excel for the Macintosh is an example of a program that can't open a file with an empty type. This is why there are two maps for Excel (the fourth and fifth maps above). The incoming map for Excel is like the one for Microsoft Word, but the outgoing map explicitly defines the Macintosh type.

The last map shows that you can have both Macintosh creator/type and MIME type/subtype in one entry. This map says that if an incoming message has an attachment that is encoded in MIME and has the "application/postscript" type/subtype, or has a BinHex attachment with the Macintosh type of "EPSF," then the resulting file will have an extension of ".eps." Similarly, if an outgoing message has an attachment with the extension ".eps" and if the MIME encoding is being used for the message, then the attachment will get the "application/postscript" MIME type/subtype. If the message was using the BinHex encoding, then the attachment would get the Macintosh type of "EPSF."

But what happens if an attachment matches more than one map? Eudora will try and find the best match. For example, if you had the following [Mappings] section:

```
[Mappings]

in=xls, XC EL, , ,

in=xlc, XCEL, xLC3, ,
```

and you received a message with an attachment that had a Macintosh creator of "XCEL" and a Macintosh type of "XLC3" (a Microsoft Excel Chart), then the file would get an extension of ".xlc" since the first map only matched the Macintosh creator, but the second map matched both the Macintosh creator and type.

Eudora can receive attachments that have both a MIME type/subtype and a Macintosh creator/type. Eudora understands attachments with the MIME type/subtype "application/apple file," which has Macintosh creator/type information embedded in it. With this type of attachment, Eudora will consider a

match with the Macintosh creator/type as a “better” match than a match with the MIME type/subtype.

Finally, if an incoming attachment matches two different maps to the same degree (e.g., both maps have the same MIME type/subtype with different file extensions), then Eudora will use the file extension in the first matching map.